# Frequency-Modulation Vowel Maps in Normal-Hearing and Hearing-Impaired Listeners

**Santurette, Sébastien; Vatti, Marianna; Pontoppidan, Niels henrik; Dau, Torsten**

*Publication date:*
2013



*Citation (APA):*
Santurette, S., Vatti, M., Pontoppidan, N. H., & Dau, T. (2013). *Frequency-Modulation Vowel Maps in Normal-Hearing and Hearing-Impaired Listeners*. Poster session presented at 36th Annual Midwinter Meeting of the Association for Research in Otolaryngology, Baltimore, MD, United States.

#223
ASSOCIATION FOR RESEARCH IN OTOLARYNGOLOGY
FOUNDED IN 1973
36th MIDWINTER RESEARCH MEETING
FEBRUARY 16-20, 2013
BALTIMORE, MD, USA

# FREQUENCY-MODULATION VOWEL MAPS IN NORMAL-HEARING AND HEARING-IMPAIRED LISTENERS

Contact: Sébastien Santurette <ses@elektro.dtu.dk>, Marianna Vatti <mav@eriksholm.com>

Sébastien Santurette[1], Marianna Vatti[1,2], Niels H. Pontoppidan[2], and Torsten Dau[1]

[1]*Centre for Applied Hearing Research, Technical University of Denmark;* [2]*Eriksholm Research Centre, Oticon A/S*

## INTRODUCTION

Sound emitted by most natural vibrating sources is not steady in pitch but contains frequency fluctuations over time during which the harmonicity of the frequency components is typically preserved, as for instance in the vibrato of human singing voices (*e.g.*, McAdams, 1984; Maher and Beauchamp, 1990). This coherent frequency modulation (CFM) has been shown to enhance the subjective fusion of the sound's frequency components (Sundberg, 1994), with implications for the ability of listeners to recognize and segregate vowels (*e.g.*, McAdams, 1989; Demany and Semal, 1990; Divenyi *et al.*, 1997). There is compelling evidence that sensorineural hearing loss is, in addition to a loss of sensitivity, often accompanied by a deterioration of frequency discrimination and of the ability to detect frequency changes over time, as reflected by elevated frequency-modulation difference limens (FMDLs) (*e.g.*, Zurek and Formby, 1981; Moore and Skrodzka, 2002). Such deterioration, arguably due to a loss of resolution in the representation of tonotopic activity, temporal phase-locked information, or both, may thus have consequences for the perception of vocal vibrato, and thereby vowel quality and identification, in hearing-impaired (HI) listeners. The present study investigated the ability of HI listeners to perceive a sung vowel based on the addition of CFM to a steady complex tone, compared to a group of normal-hearing (NH) listeners. This was achieved by obtaining "vowel maps" in the two groups as a function of the two primary acoustic parameters of vocal vibrato: FM rate and FM excursion.

## RESEARCH QUESTIONS

- What is the extent of the vowel map in NH listeners along the FM-rate and FM-excursion dimensions?
- Is such a map affected by hearing impairment, and if so, along which dimensions?

## METHODS

### Stimulus configuration

- Harmonic complex tone with first 8 harmonics of vowel /oh/
- CFM applied by adding the same frequency shift M (cents) to all N components:

$$y(t)=\sum_{k=1}^{N} A_k \cos\left(2\pi f_k t + 2\pi \int_0^t f_k^{\frac{c_m(\tau)}{1200}}\, d\tau\right) \quad , \quad c_m(\tau)=M\sin(2\pi f_m \tau)$$

($k$: harmonic number, $f_m$: FM rate, $A_k$: harmonic amplitude, $f_k$: harmonic frequency)

- Shimmer and jitter added for better simulation of natural vocal vibrato
- Three temporal segments with "old+new heuristic" (Bregman and Ahad, 1996) such that adding CFM leads to the fusion of all components into a singing voice [Fig. 1]

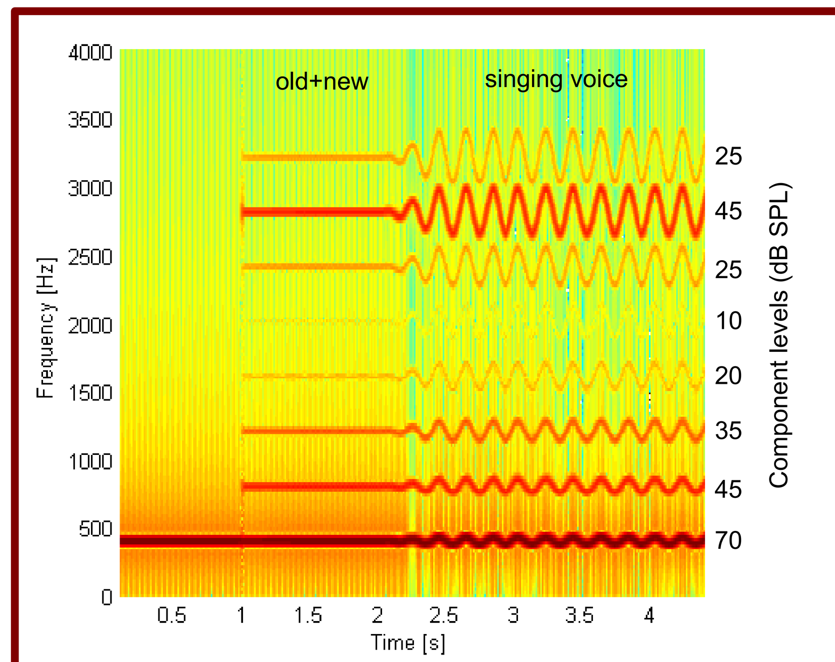


**Figure 1**: Spectrogram of the synthesized-vowel stimulus. The first temporal segment contains the fundamental component at 400 Hz, the second segment all eight harmonics, and the third segment all harmonics with applied CFM.

In the second segment, the fundamental is segregated from the added harmonics.

In the third segment, the addition of CFM creates fusion of the fundamental with the other harmonics and a singing voice emerges.

### Procedure

- Tracking of the "sweet spot" area for which a singing voice emerges in the third stimulus segment as a function of FM rate and FM excursion [Fig. 2]
- One FM parameter was kept constant while the other was adjusted in a 1-interval, 2-AFC yes/no task with a 1-up 1-down paradigm
- Thresholds were approached from outside the sweet spot (unnatural vibrato) at FM rates of [3, 4, 5, 6, 7, 8] Hz and FM excursions of [21, 35, 49, 63, 77, 91] cents
- 6 reversals, step sizes of [0.5, 0.14] Hz (FM rate) and [14, 4] cents (FM excursion)
- 3 repetitions per subject for each of the 24 conditions

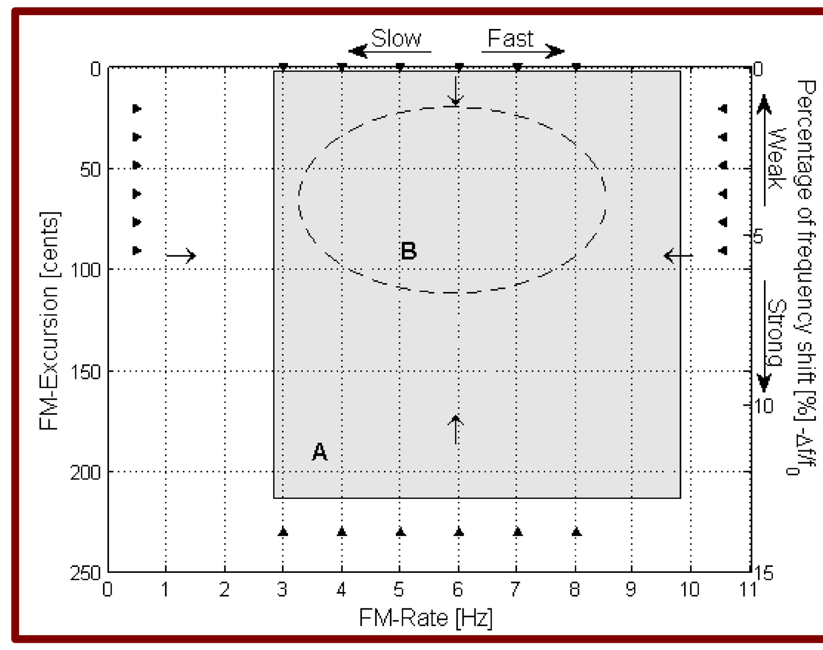


**Figure 2**: Hypothetical vowel map along the FM-rate and FM-excursion dimensions.

A: FM range of human vocal vibrato as reported in the literature.

B: Hypothetical "sweet spot" to be determined in this study. Filled triangles indicate reference values used in the experimental procedure. Arrows indicate the four directions from which the different boundaries were approached.

### Subjects

- 14 NH listeners (8 musicians [NHm], 6 non-musicians [NHo])
- 12 HI listeners (7 musicians [HIm], 5 non-musicians [Hio])
  - Flat or sloping sensorineural hearing loss [Table 1]

**Table 1**: Individual listener characteristics for the vowel-mapping experiment. For each listener, the following details are given:

- Age in years
- Low-frequency (0.125–1 kHz) average pure-tone audiogram in dB HL (LF-HL)
- High-frequency (1.5–6 kHz) average pure-tone audiogram in dB HL (HF-HL)
- Equivalent rectangular bandwidth (ERB) of the auditory filter at 400 Hz in the left ear in Hz, estimated with the notched-noise method

All listeners had nearly symmetric audiograms in the left and right ears, except Him5 who had a unilateral hearing loss in the left ear.

Absolute hearing thresholds are thus averaged between ears, except for subject Him5 for whom the left-ear thresholds are given.

All musicians had at least 5 years of musical training and practiced an instrument or singing on a regular basis.

Non-musicians did not receive any form of musical training and did not practice any instrument or sing.

| Subject | Age | LF-HL | HF-HL | ERB |
|---|---|---|---|---|
| NHm1 | 26 | -3.5 | -2.5 | 92.3 |
| NHm2 | 33 | -4.0 | 3.0 | |
| NHm3 | 27 | 1.0 | 3.0 | 70.9 |
| NHm4 | 28 | 0.0 | -2.5 | 66.4 |
| NHm5 | 27 | 0.0 | 3.0 | 85.0 |
| NHm6 | 29 | -1.0 | 0.5 | 93.5 |
| NHm7 | 67 | 7.5 | 17.0 | 77.9 |
| NHm8 | 71 | 4.5 | 18.5 | 110.0 |
| NHo1 | 27 | | | 75.8 |
| NHo2 | 23 | 3.1 | 1.9 | 95.5 |
| NHo3 | 24 | -2.5 | 0.0 | 106.5 |
| NHo4 | 23 | 6.3 | 0.6 | 94.2 |
| NHo5 | 24 | 3.1 | 13.1 | 83.6 |
| NHo6 | 24 | 0.0 | 1.3 | 82.2 |
| HIm1 | 25 | 39.5 | 51.5 | 208.3 |
| HIm2 | 79 | 16.0 | 48.5 | 94.8 |
| HIm3 | 71 | 14.5 | 43.0 | 102.5 |
| HIm4 | 68 | 15.0 | 56.5 | |
| HIm5 | 68 | 28.0 | 41.0 | 106.2 |
| HIm6 | 73 | 12.5 | 35.0 | 72.0 |
| HIm7 | 77 | 21.0 | 21.0 | 96.3 |
| HIo1 | 72 | 54.0 | 58.5 | |
| HIo2 | 70 | 45.0 | 52.0 | 125.3 |
| HIo3 | 69 | 48.0 | 46.0 | 161.6 |
| HIo4 | 63 | 41.0 | 46.0 | 78.7 |
| HIo5 | 77 | 45.5 | 61.0 | 239.9 |

### Presentation level

- Stimuli were presented diotically at a moderate sound level
- Component levels for NH listeners as in Bregman and Ahad (1996) [Fig. 1]
- Levels were adjusted for equal loudness in HI listeners with categorical loudness scaling such that the loudness of each component corresponded to the average loudness of that component in the NH group

## RESULTS

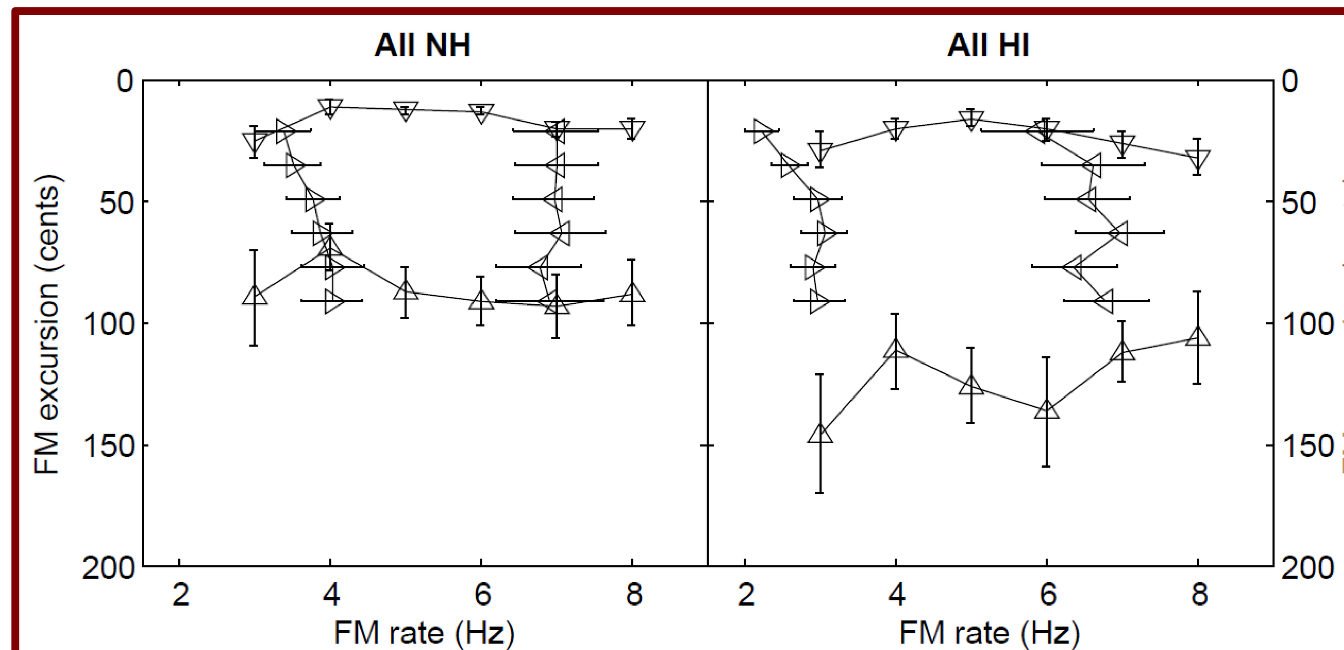


**Figure 3**: Average frequency-modulation vowel maps for NH and HI listeners. The mean and standard error over all subjects within each group is given for each threshold.

### Effects of hearing impairment and musical experience

- Broader sweet spots in HI listeners, shifted towards higher FM excursions and lower FM rates compared to NH listeners [Figs. 3, 4, 5]
- 2-way ANOVA on mean FM rates and square-root transformed mean FM excursions for each boundary, with hearing loss and musical experience as factors [Fig. 5]:
  - Significant effect of hearing loss for all sweet-spot boundaries except upper FM rate
  - Significant effect of musical experience for lower FM excursion boundary only
- No correlation of mean FM-parameter thresholds with HF-HL or ERB for any boundary
- Correlation of mean thresholds with LF-HL for lower FM excursion boundary only ($\rho$=0.88)
  - Large individual differences [Fig. 4] not explained by audibility or frequency selectivity at F0

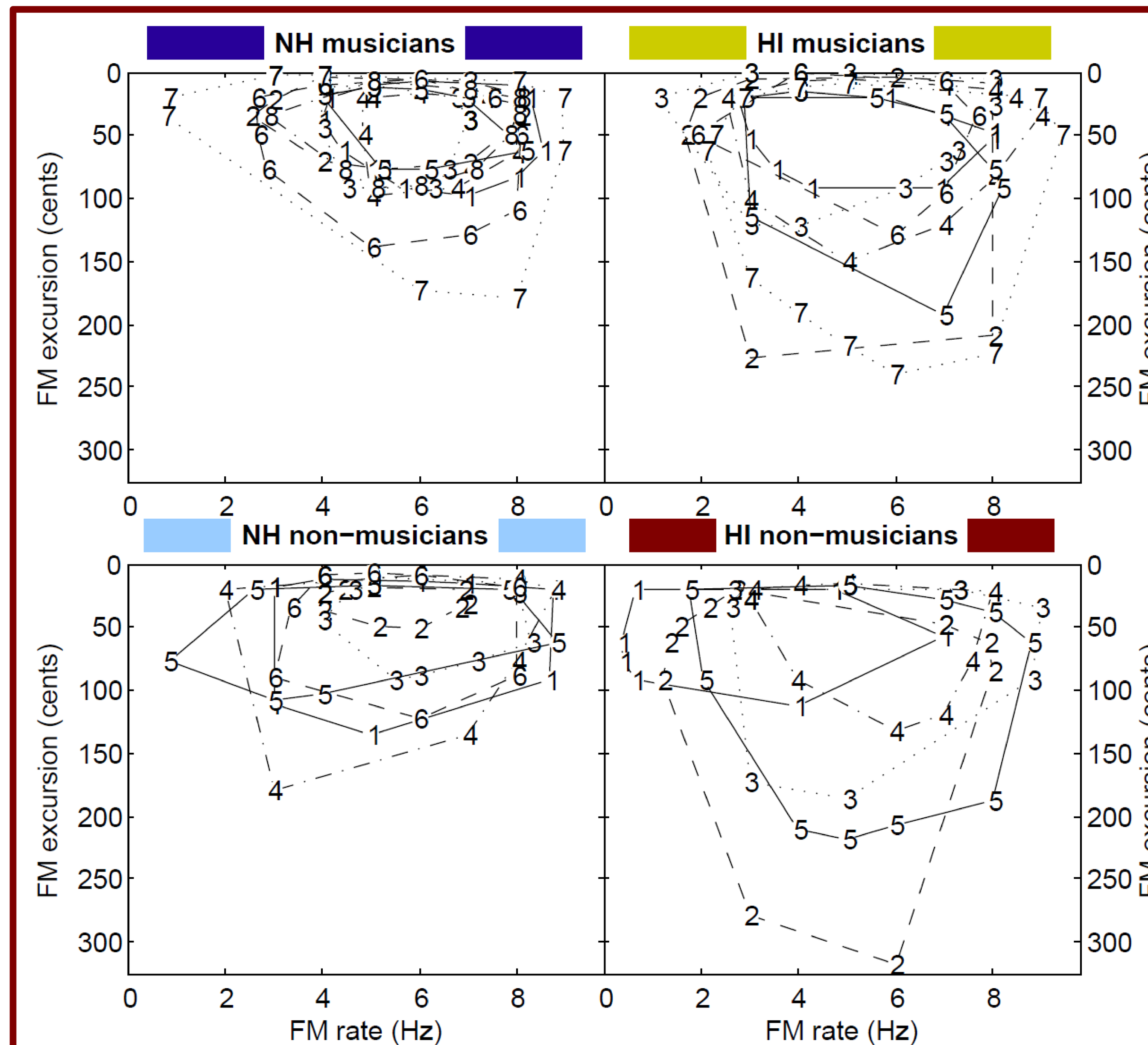


**Figure 4**: Individual sweet spots plotted as convex hulls. The plots are grouped according to hearing group (NH vs. HI) and musical experience (musicians vs. non-musicians).

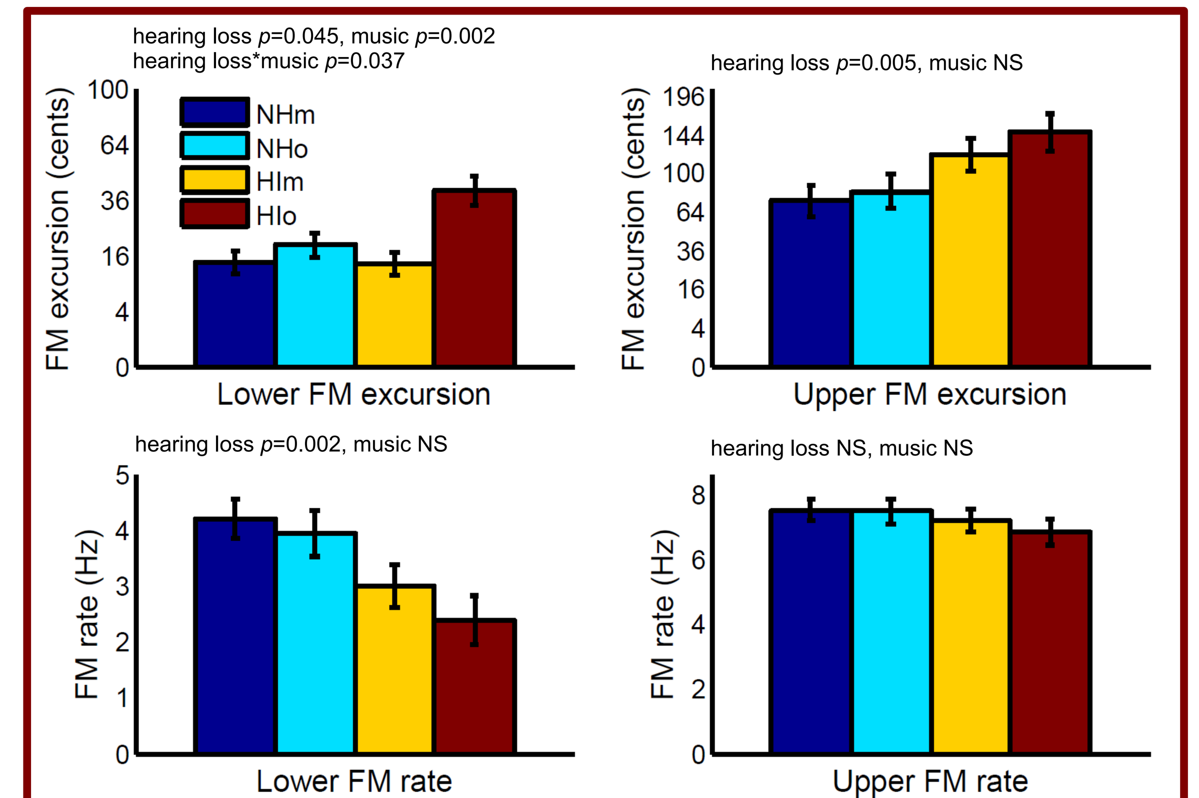


**Figure 5**: Marginal means and standard errors for each of the four sweet-spot boundaries as a function of hearing group and musical experience.

### Potential explanations for individual differences in the sweet-spot area

- The HI listeners' preference for larger FM excursions may stem from the need for faster rates of instantaneous frequency changes (Horii, 1989) than NH listeners in order to perceive the same amount of vibrato.
- Alternatively, HI listeners may rely more on peak excursions in the FM cycle.
- Potential difficulties of HI listeners with FM-rate and FM-excursion discrimination would also be expected to lead to less well-defined sweet spots.
- Auditory filter bandwidths around higher harmonics may play a more crucial role than frequency selectivity around the vowel F0.
- Individual listeners may have different criteria for "sweetness".
- The chosen loudness compensation strategy for HI listeners may have introduced timbre differences between subjects.

## CONCLUSIONS

- In NH listeners, adding CFM to an unmodulated complex tone was sufficient to evoke the perception of a singing voice for FM rates between 4.1 and 7.5 Hz and FM excursions between 17 and 83 cents on average.
  - These values may provide some guidelines when constructing synthetic-vocal stimuli for which a realistic sung vibrato is desired.
- Hearing loss was found to affect the perception of a sung vowel based on FM-rate and FM-excursion cues.
  - Further work is needed to clarify the role of deficits in detection or discrimination of FM parameters, temporal fine-structure processing, and the ability to follow the rate of frequency changes, for the perception of vocal vibrato and vowel quality in HI listeners.

## REFERENCES

Demany, L., & Semal, C. (1990). *Perception and Psychophysics* 48:436–444.
Divenyi, P. L., Carré, R., & Algazi, A. P. (1997). *IEEE Mohonk Mountain Workshop on Applications of Signal Processing to Audio and Acoustics*.
Horii, Y. (1989). *J. Speech Hear. Res.* 32:829–836.
Maher, R., & Beauchamp, J. (1990). *Applied Acoustics* 30:219–245.
McAdams, S. (1984). PhD thesis, Stanford University.
McAdams, S. E. (1989). *J. Acoust. Soc. Am.*, 86:2148–2159.
Sundberg, J. (1994). *Speech, Music and Hearing Quarterly Progress and Status Report (STL-QPSR)* 35:45–68.
Moore, B. C. J., & Skrodzka, E. (2002). *J. Acoust. Soc. Am.* 111:327–335.
Zurek, P. M., & Formby, C. (1981). *J. Speech Hear. Res.* 24:108–112.